

PARTIAL TRANSLATION OF JP 11 (1999)-85988 A

Publication Date: March 30, 1999

Title of the Invention: FACE IMAGE RECOGNITION SYSTEM

Patent Application Number: 9-239307

Filing Date: September 4, 1997

Inventors: Naoki Sashita et al.

Applicant: FUJITSU LTD.

(Page 5, right column, line 40 – page 6, left column, line 2)

[0060] Figure 11 is a flow chart showing processing contents of the recognizing process part 1022 in learning phase. As shown in Figure 11, in the face image recognition system of the present invention, prior to the selection of a window based on the edge intensity in accordance with the Eigen-Window method described above, selection of a window is conducted based on skin color information (S1101). This is one improvement of the Eigen-Window method based on that a recognition target is a human's face picture.

[0061] The selection of a window based on skin color information refers to processing in which, in initial setting processing, a window is selected only from a region corresponding to a skin color in an learning image, based on skin color information recognized by the face image recognition system of the present embodiment. Because of such processing, the efficiency of the subsequent processing, such as the selection of a window based on the edge intensity, can be enhanced.

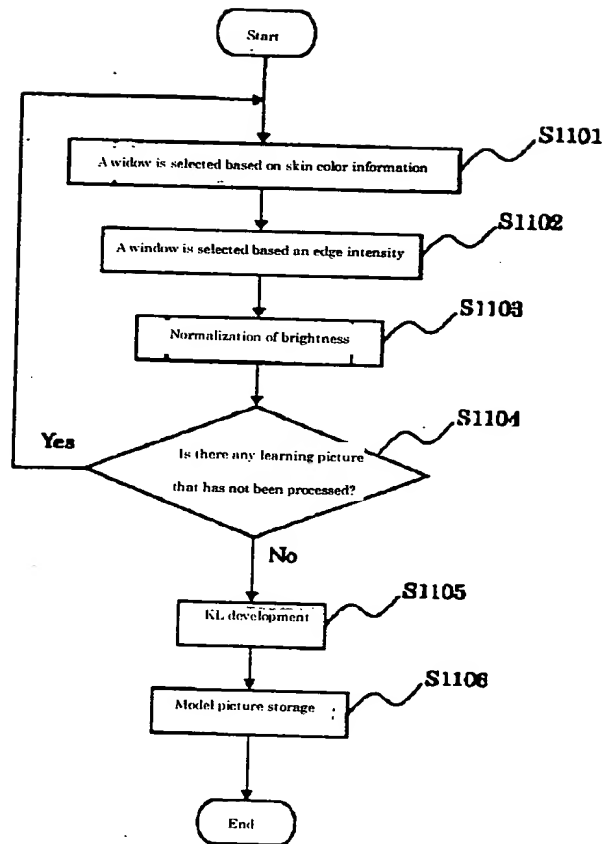
(Page 6, left column, lines 30 – 44)

[0066] When the above-mentioned selection of a window based on skin color information is completed, the process returns to the flow chart in Figure 11, and a window is selected based on the edge intensity (S1102). The selection of a window based on the edge intensity has already been described in the above-mentioned description of the Eigen-Window method. Therefore, the detailed description

thereof will be omitted here. In the face image recognition system of the present embodiment, the edge intensity is calculated based on a gradation value of an image only with respect to a pixel representing the selected skin color in the processing of the selection of a window based on skin color information, and a window is selected.

[0067] Next, in the window image recognition system of the present embodiment, normalization processing of brightness is conducted (S1103). This is one improvement of the Eigen-Window method, for realizing robust recognition to some degree with respect to changes in appearance of a target caused by changes in illumination and environment that may occur during capturing of a face image.

Fig. 11





## PATENT ABSTRACTS OF JAPAN

(11) Publication number: **11085988 A**(43) Date of publication of application: **30 . 03 . 99**

(51) Int. Cl.

**G06T 7/00**  
**G06T 1/00**
(21) Application number: **09239307**(22) Date of filing: **04 . 09 . 97**(71) Applicant: **FUJITSU LTD**
(72) Inventor: **SASHITA NAOKI**  
**MASUMOTO DAIKI**  
**OSADA SHIGEMI**
(54) **FACE IMAGE RECOGNITION SYSTEM**

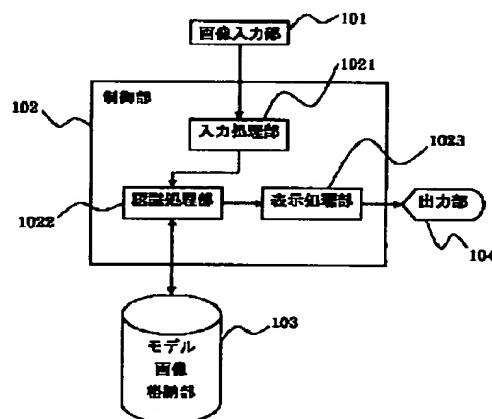
## (57) Abstract:

**PROBLEM TO BE SOLVED:** To cope with a position shift, hiding, background noise or the like, in an image of a face pattern by selecting a features small area from face image information that a face image information input means inputs and converting information included in respective selected small areas into a spot set projected on a characteristic space.

**SOLUTION:** The face image information is inputted through an image input part 101. A recognizing process part 1022 extracts feature data from the image information that an input process part 1021 receives in learning phase and stores them as model image information in a model image storage part 103. Further, the recognizing process part 1022 extracts feature data from the face image information of a recognition object accepted by the input process part 1021 in recognition phase and matches them against the information stored in the model image storage part 103 in the learning phase to recognize the face image. The result of the recognition is sent to a display process part 1023. In this case, small featured areas are selected from the face image information and information included in the

respective small areas is converted into a spot set projected on the regular space.

COPYRIGHT: (C)1999,JPO



(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平11-85988

(43)公開日 平成11年(1999) 3 月30日

(51)IntCl.<sup>4</sup>

G 0 6 T 7/00  
1/00

識別記号

F I

G 0 6 F 15/62 4 6 5 K  
15/66 3 3 0 R  
15/70 4 6 0 A  
4 6 5 A

審査請求 未請求 請求項の数5 O L (全 10 頁)

(21)出願番号 特願平9-239307

(22)出願日 平成9年(1997) 9 月 4 日

(71)出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番  
1号

(72)発明者 指田 直毅

神奈川県川崎市中原区上小田中4丁目1番  
1号 富士通株式会社内

(72)発明者 増本 大器

神奈川県川崎市中原区上小田中4丁目1番  
1号 富士通株式会社内

(72)発明者 長田 茂美

神奈川県川崎市中原区上小田中4丁目1番  
1号 富士通株式会社内

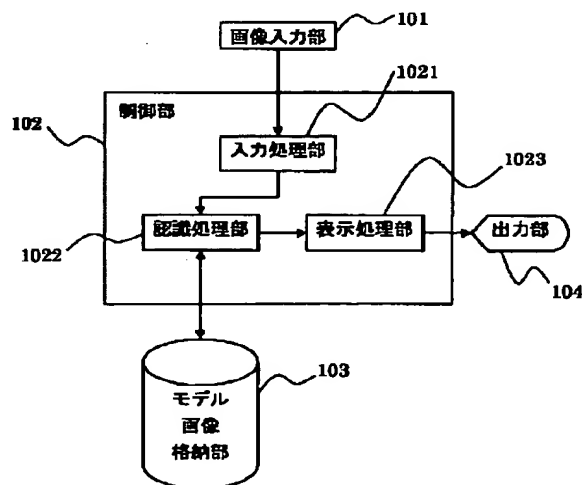
(74)代理人 弁理士 井桁 貞一

(54)【発明の名称】 顔画像認識システム

(57)【要約】

【課題】 認識しようとする画像に含まれる認識対象の位置が学習時と比べてずれている場合、認識対象の一部が隠れている場合、背景ノイズが存在する場合、複数の認識対象が存在する場合などにも対応できる顔画像識別システムを提供すること。

【解決手段】 入力処理部が取り込んだ顔画像情報を、モデル画像格納部に格納されているモデル画像と照合して顔画像の認識を行う際に、認識処理部において、従来法の欠点を克服したEigen-Window法に基づく画像認識手法を適用した処理を行う。



## 【特許請求の範囲】

【請求項 1】 顔画像情報を取り込む顔画像情報入力手段と、前記顔画像情報入力手段が取り込んだ顔画像情報を変換する顔画像情報変換手段と、前記顔画像情報変換手段により変換された情報を記憶するモデル画像記憶手段と、前記顔画像情報変換手段により変換された情報と、前記モデル画像記憶手段に記憶されている情報とを照合する照合手段とを備える顔画像認識システムにおいて、前記顔画像情報変換手段は、前記顔画像情報入力手段が取り込んだ顔画像情報から特徴的な小領域を選択し、選択された各小領域に含まれる情報を固有空間に投影した点集合に変換することを特徴とする顔画像認識システム。

【請求項 2】 前記モデル画像記憶手段は、人物の顔を複数の視点から撮影した複数の顔画像情報について、それぞれ前記顔画像情報変換手段により変換した情報を記憶する請求項 1 記載の顔画像認識システム。

【請求項 3】 前記モデル画像記憶手段は、複数の人物について、前記変換した情報を記憶する請求項 2 記載の顔画像認識システム。

【請求項 4】 前記顔画像情報変換手段はさらに、顔画像情報から人間の顔の部分に該当する情報を抽出する抽出手段を有し、前記顔画像情報入力手段がとり込んだ顔画像情報から特徴的な小領域を選択する際に、前記抽出手段により抽出された情報から前記小領域を選択する請求項 1 記載の顔画像認識システム。

【請求項 5】 前記顔画像情報変換手段は、前記特徴的な小領域を選択した後、当該選択された小領域の画素値を要素とする窓画像ベクトルに対して、明るさの正規化処理を施す請求項 1 又は 4 記載の顔画像認識システム。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、人物の顔画像情報を変換した情報をデータベースに蓄積しておき、認識時には、入力された顔画像情報に所定の変換を施した後、前記データベースに記憶された情報と照合することにより、顔画像を認識する顔画像認識システムに関する。

## 【0002】

【従来の技術】顔が表す情報をコンピュータで読み取ることができるになれば、人間と機械との自然なインターフェースの実現に役立つとの期待から、近年、コンピュータで顔画像を認識しようとする試みが活発に行われている。また、顔画像に基づく個人識別についても、コンピュータでこれを代替、支援できるようになれば、犯罪捜査などのための顔画像データベース認識システムや、セキュリティのためのアクセスコントロールシステムなどへの応用が期待できる。このように、顔画像認識技術は、マンマシンインターフェースの高度化に際しての重要な課題の一つであり、今後ますますその重要性を

増していくものと予想される。

【0003】コンピュータによる顔画像認識においては、顔の姿勢や、撮影位置の照明に代表される撮影条件などに伴い、見え方が変化する入力画像のボタンに対して、例えば、その人物の氏名の如く、識別結果の情報を安定に出力することが求められる。

【0004】このような問題に対して、画像中の顔が正面を向いているという前提のもとに、入力画像中から目や口などの部品（以下、「顔部品」という。）を正確に抽出し、これら顔部品の形状や配置における個性をもとにして識別を行おうとする、顔部品ベースの手法がある。

【0005】しかし、この顔部品ベースの手法では、そもそも現実の環境で撮影された顔画像から、顔部品の線図的な形状やその位置などを正確に抽出すること自体大変困難な問題であり、さらに構造的に類似した個人の微妙な顔パタンの差異を線図的な形状だけで記述することにも限界がある。

【0006】そこで、最近になって注目を集めているアプローチとして、顔画像のボタンを各画素における濃淡値の二次元配列として表現し、そのマッチングによって識別を行おうとするボタン整合の手法が提案されている。この手法の代表的なものとしては、高次元の画像データを圧縮した固有空間と呼ばれる低次元空間上でボタン照合を行う固有空間法（例えば、Hiroshi Murase and Shree K.Nayer, "Visual Learning and Recognition of 3-D Objects from Appearance", International Journal of Computer Vision, Vol.14, No.1, pp.5-24 (1995) に開示されている）等がある。

【0007】この固有空間法に基づく顔画像認識システム（例えば、特開平 5 - 2 0 4 4 2 号公報に開示されている。）は、ほぼ一定の照明条件のもとで撮影され、かつ正確な位置合わせ、および顔サイズの正規化処理が既に施されている正面顔の識別に対して、有効であることが示されている。

## 【0008】

【発明が解決しようとする課題】しかしながら、上記従来の固有空間法においては、入力画像中の顔パタンの正確な位置決めと正規化処理を必要とするという点からも分かるように、画像中に含まれる顔パターンの画像内での位置ずれや、隠れ、背景ノイズなどに対して影響を受けやすいという問題点を有していた。

【0009】また、上記従来の固有空間法を用いた顔画像認識システムでは、正面を向いている顔画像であり、かつ顔パタンの正確な位置決めと正規化が行われているという制約条件の下では良好な結果を示すものの、現実的な環境で撮影した場合に起こることが予想される、顔の位置や姿勢（あるいは観測の視点）、照明条件、背景条件、時間経過などによる見え方の変化に対してロバストな顔画像認識を実現することに関しては、依然未解決

の問題として残されている。

【0010】本発明は、認識しようとする画像に含まれる認識対象の位置が学習時と比べてずれている場合、認識対象の一部が隠れている場合、背景ノイズが存在する場合、複数の認識対象が存在する場合などにも対応できる顔画像識別システムを提供することを目的とする。

【0011】

【課題を解決するための手段】前記目的を達成するため、本発明の顔画像認識システムは、顔画像情報を取り込む顔画像情報入力手段と、前記顔画像情報入力手段が取り込んだ顔画像情報を変換する顔画像情報変換手段と、前記顔画像情報変換手段により変換された情報を記憶するモデル画像記憶手段と、前記顔画像情報変換手段により変換された情報と、前記モデル画像記憶手段に記憶されている情報とを照合する照合手段とを備える顔画像認識システムにおいて、前記顔画像情報変換手段は、前記顔画像情報入力手段が取り込んだ顔画像情報から特徴的な小領域を選択し、選択された各小領域に含まれる情報を固有空間に投影した点集合に変換することを特徴とする。

【0012】また、前記モデル画像記憶手段は、人物の顔を複数の視点から撮影した複数の顔画像情報について、それぞれ前記顔画像情報変換手段により変換した情報を記憶するのが好ましい。当該情報を記憶しておき、入力された認識対象の顔画像と照合することにより、認識対象の顔画像が正面から撮影されたものでない場合でも、顔画像の認識を行うことが可能となるからである。

【0013】また、前記モデル画像記憶手段は、複数の人物について、前記変換した情報を記憶することが好ましい。当該情報を記憶しておき、入力された認識対象の顔画像と照合することにより、認識対象である人物を特定することが可能となるからである。

【0014】また、前記顔画像情報入力手段として、ビデオカメラを用いることができる。この場合には、容易に認識対象である顔画像を撮影して、本発明の顔画像認識システムに取り込むことができる。

【0015】また、前記顔画像情報入力手段として、イメージスキャナを用いることもできる。この場合には、既に撮影されている顔画像の認識を行うことが可能となる。また、前記顔画像情報入力手段として、外部記憶装置を用いることもできる。近年、デジタルカメラのように、顔画像情報をデジタル情報として処理する装置も多く用いられるようになっており、前記顔画像情報入力手段として、外部記憶装置を用いることにより、デジタル化された情報についても、本発明の顔画像認識システムに取り込むことが可能となる。

【0016】前記顔画像情報変換手段はさらに、顔画像情報から人間の顔の部分に該当する情報を抽出する抽出手段を有し、前記顔画像情報入力手段がとり込んだ顔画像情報から特徴的な小領域を選択する際に、前記抽出手

段により抽出された情報から前記小領域を選択するのが好ましい。顔画像情報の処理においては、一般に取り扱うべき情報量が多く、処理に時間がかかる場合が多いが、前記抽出手段を備えることにより、処理量の軽減を図ることができるからである。

【0017】また、前記顔画像情報変換手段は、前記特徴的な小領域を選択した後、当該選択された小領域の画素値を要素とする窓画像ベクトルに対して、明るさの正規化処理を施すことが好ましい。かかる処理を行うことにより、撮影時の照明の条件等による明るさの変化に対してもロバストな認識を行うことが可能となるからである。

【0018】以上のように、本発明の顔画像認識システムを用いることにより、認識しようとする画像に含まれる認識対象の位置が学習時と比べてずれている場合、認識対象の一部が隠れている場合、背景ノイズが存在する場合、複数の認識対象が存在する場合などにも対応できる顔画像識別システムを提供することができる。

【0019】

【発明の実施の形態】以下、本発明の実施の形態について、図面を参照しながら説明する。

【0020】図1は、本発明の顔画像認識システムの全体構成を示すブロック図である。同図に示されるように、本発明の顔画像認識システムは、顔画像情報入力手段としての画像入力部101、顔画像情報変換手段、及び照合手段としての制御部102、モデル画像記憶手段としてのモデル画像格納部103、出力部104を備えている。

【0021】画像入力部101としては、例えばビデオカメラ、イメージスキャナ等が用いられ、画像入力部101を介して、学習画像及び認識対象画像となる顔画像情報が入力される。また、画像入力部101としては、補助記憶装置を接続して使用することも可能である。

【0022】制御部102は、入力処理部1021、認識処理部1022、表示処理部1023を備える。入力処理部1021は、画像入力部101からの入力を受け付ける。

【0023】認識処理部1022は、後述する学習フェーズでは、入力処理部1021が受け付けた画像情報から特徴データを抽出し、モデル画像情報として、モデル画像格納部103に格納する。また、認識処理部1022は、後述する認識フェーズでは、入力処理部1021が受け付けた認識対象の顔画像情報から特徴データを抽出し、学習フェーズでモデル画像格納部103に格納された情報と照合することにより顔画像の認識を行い、認識の結果を表示処理部1023に送る。特徴データの抽出処理については、後に詳細に説明する。

【0024】表示処理部1023は、認識処理部1022による認識の結果を出力部104に出力する。

【0025】モデル画像格納部103には、前記認識処

理部1022により生成されたモデル画像情報が格納される。

【0026】出力部104には、顔画像の認識結果等が出力される。

【0027】ここで、本発明に係る顔画像認識システムで用いるEigen-Window法の一般的な概念について説明する。Eigen-Window法は、上記従来の固有空間法を改良したものであり、固有空間法が、認識対象が含まれる画像全体を一つのモデルとして扱うのに対して、画像中の認識対象を特徴的な小領域（以下、「窓画像」という。）に分割し、これらの集合で認識対象をモデル化する手法である。Eigen-Window法を用いる三次元物体の認識では、認識対象を部分的な要素の集合として捉えることにより、固有空間法では実現できなかった、画像中に含まれる認識対象の位置ずれや隠れ、背景ノイズ、画像中に認識対象が複数存在する場合などに対応した認識が可能となる。

【0028】以下、Eigen-Window法に基づく一般的な三次元物体の認識の処理過程について説明する。

【0029】Eigen-Window法の処理過程は、入力画像からモデル画像情報を生成する学習フェーズ、学習フェーズで作成されたモデル画像情報と認識対象画像とを照合\*

$$Z_m = [z_1^1, z_1^2, \dots, z_m^1, \dots, z_m^{N_m}]$$

【0035】

【数2】

$$Q = (Z_m - c)(Z_m - c)^T \in \mathbb{R}^{N^2 \times N^2}$$

【0036】ここで、 $c$ は選択された全ての窓画像の平均濃淡値ベクトルを表す。この行列 $Q$ の固有ベクトル $e_i$ を対応する固有値 $\lambda_i$ の順にソートし、ある閾値以上の固有値を持つ固有ベクトルを基に、窓画像が持つ元々の次元 $N^2$ に対して十分小さな次元 $k$ で構成される固有空間への変換行列 $R$ を構成する。変換行列 $R$ は、具体的には下記の（数3）で表される。

【0037】

【数3】

$$R = [e_1, e_2, \dots, e_k] \quad (k \ll N^2)$$

【0038】変換行列 $R$ によって張られる $k$ 次元固有空間へ、学習窓画像 $Z_m$ を投影し、対応する点群行列 $G_m$ を導く（これら一連の変換処理をKarhunen-Loeve展開、以下、略して「KL展開」という。）。点群行列 $G_m$ は、具体的には、下記の（数4）で表される。

【0039】

【数4】

$$G_m = [g_1^1, \dots, g_m^1, \dots, g_m^{N_m}] = R^T(Z_m - c)$$

【0040】導かれた点群行列 $G_m$ （以下、 $G_m$ を、「学習窓画像対応点」という。）を表す情報が、モデル画像格納部103に格納され、学習フェーズが終了す

\*することにより、認識対象画像の認識を行う認識フェーズの2つのフェーズから構成される。

【0030】図2は、学習フェーズにおける処理過程の概略を示す図である。同図に示されるように、学習フェーズでは、まず、一つの認識対象を多視点から撮影した一連の学習画像（ $M$ 枚）を取り込み、まず、エッジ強度に基づく窓選択を行う。

【0031】エッジ強度に基づく窓選択とは、画素の濃淡値を基にエッジ強度を計算し、ある閾値以上のエッジ強度を持つ画素を中心とする $N$ 画素 $\times N$ 画素の矩形小領域を窓画像として選択する処理である。

【0032】図3は、窓選択について説明するための図である。学習画像から $I$ 個の窓画像が選択されたとする、以下の説明では、 $m$ 枚目の学習画像から選択された $i$ 番目の窓画像を $Z_m^i$ と表すことにする。

【0033】エッジ強度に基づく窓選択処理が終了すると、 $m$ 枚目の学習画像に含まれる $i$ 番目の窓画像の濃淡値を要素とする列ベクトル $z_m^i$ から構成される、下記（数1）に示される行列 $Z_m$ に対して、下記（数2）に示す共分散行列 $Q$ を計算する。

【0034】

【数1】

$$(1 \leq m \leq M)$$

る。図4は、 $k=3$ の場合の固有空間における学習窓画像対応点 $G_m$ の様子を表す図である。

【0041】次に、認識フェーズにおける処理について説明する。図5は、認識フェーズにおける処理内容の概要を示す図である。同図に示されるように、認識フェーズにおいては、認識対象となる画像の入力を受けて、入力された認識対象画像に対して、学習フェーズと同様に、エッジ強度に基づく窓選択およびKL展開を行う。認識フェーズにおけるKL展開では、選択された入力窓画像行列 $Z_{in}$ に対応する固有空間上の点群行列 $G_{in}$ が求められる。 $G_{in}$ は、具体的には、下記の（数5）によって表される（以下、 $G_{in}$ を、「入力窓画像対応点」という）。これは、認識対象画像に関して上記（数4）と同様の内容を表すものである。

【0042】

【数5】

$$G_{in} = R^T(Z_{in} - c)$$

【0043】上記の処理において得られた入力窓画像対応点 $G_{in}$ に含まれる各入力窓画像対応点 $g_{in}^i$ について、学習フェーズでモデル画像格納部103に格納された $G_m$ を表す情報との照合処理を行うことによって、識別対象画像が示す物体の特定、および当該物体の姿勢の特定が行われる。

【0044】以下、認識フェーズにおける照合処理の内容について説明する。照合処理は、最近傍点の探索と、



投票とからなる。

【0045】最近傍点の探索処理においては、固有空間上のベクトル距離を評価基準として、各入力窓画像対応点  $g_{i,n}^{'}$  に対して、対応点群行列  $G_{\cdot}$  の中から最小距離を示す学習窓画像対応点  $g_{m,n}^{'}$  を探索し、各入力窓画像  $z_{i,n}^{'}$  に見えが最も類似する学習窓画像  $z_{m,n}^{'}$  を求める。図6は、この際の固有空間の様子を具体的に示す図である。同図において、 $i$  番目の入力窓画像対応点  $g_{i,n}^{'}$  に対しては、学習窓画像対応点の一つ  $g_{m,n}^{'}$  とが、最も類似することを示している。この場合には、学習窓画像対応点  $g_{m,n}^{'}$  に対応する学習窓画像  $z_{m,n}^{'}$  が、入力窓画像  $z_{i,n}^{'}$  に見えが最も類似する学習窓画像として求められる。

【0046】投票処理とは、各入力窓画像  $z_{i,n}^{'}$  に対して、上記で求められた最も類似する学習窓画像  $z_{m,n}^{'}$  との画像内相対位置  $\Delta x(z_{i,n}^{'}, z_{m,n}^{'})$  を基に、 $m$  枚目の学習画像に対応する投票空間に投票する処理である。上記画像内相対位置は、具体的には、下記の(数6)により求められる。

【0047】

【数6】

$$\Delta x = p[z_{i,n}^{'}] - p[z_{m,n}^{'}]$$

【0048】上記(数6)において、 $p[\cdot]$  は、全体画像中の窓画像位置を算出する演算式である。投票が終了すると、 $m$  枚の各学習画像に対応する  $M$  個の投票空間上における得票数のピーク位置を検出することにより、入力画像中に含まれる対象に最も類似する学習画像を特定することができる。図7は、投票空間の様子を具体的に示す図である。ここで、得票数のピーク位置が複数検出された場合には、モデル画像に合致する認識対象が複数存在するものと認識することができる。

【0049】特定された学習画像に基づいて、入力画像中に含まれる認識対象の物体の種類、位置、姿勢等を推定することが可能となる。

【0050】次に、以上に説明したような概念に基づく処理により実現されるEigen-Window法の手法を、顔画像認識システムに適用する方法について説明する。

【0051】本発明に係る顔画像認識システムでは、上記に説明したようなEigen-Window法を実現するための種々の計算を、認識処理部1022が行うが、認識対象画像が人間の顔画像であることから、種々の改良を行っている。

【0052】まず、改良点の一つとして、学習フェーズに先立って初期設定処理を行うこととしている。初期設定処理とは、本実施の形態の顔画像認識システムに対して、顔画像の肌色を認識させることにより、より顔画像の認識精度を高めるとともに、認識処理の効率を向上させるべく行う処理である。

【0053】図8は、初期設定処理における認識処理部1022の処理内容を示すフローチャートである。同図

に示されるように、認識処理部1022は、まず、入力部101を介してサンプル画像読み取り処理を行う(S801)。図9は、サンプル画像の一例を示す図である。サンプル画像とは、肌色を含む画像であり、図9に示されるように、画像の中で肌色の部分には、指定領域として、利用者によってあらかじめマークが付けられている。認識処理部1022は、当該マークが付けられている肌色に相当する画素に対して、HSI変換を行う(S802)。

【0054】HSI変換とは、各画素にRed、Green、Blueの値(通常、各256段階)が記述されたRGB画像を、Hue(色相)、Saturation(彩度)、Intensity(輝度)の三属性で表されたHSI画像に変換する処理である。上記HSIの三属性は、照明条件の変化などに起因する明るさの変化に対して変動を受け難いという特性を有するため、明るさの変化に対してロバストな顔画像の認識を行うことが可能となる。

【0055】さらに、認識処理部1022は、肌色領域の抽出を行う(S803)。

【0056】肌色領域の抽出とは、前述の如くHSI変換された肌色部分の画像の持つHSI値を基に、各画素をHSI空間上の点としてプロットし、これらの点が適切に含まれるように、以下の(数7)に示すような肌色領域のパラメータHmin、Hmax、Smin、Smax、Imin、Imaxを決定する処理である。

【0057】

【数7】

$$\begin{aligned} Hmin < h < Hmax \\ Smin < s < Smax \\ Imin < i < Imax \end{aligned}$$

【0058】図10は、上記HSI空間における肌色領域の抽出について、具体的に説明するための図である。同図に示される直方体の部分が、HSI空間における肌色領域である。

【0059】以上のような初期設定処理を終了した本実施の形態の顔画像認識システムによる、学習フェーズと認識フェーズにおける認識処理部1022の処理内容について、以下に図面を参照しながら説明する。

【0060】図11は、学習フェーズにおける認識処理部1022の処理内容を示すフローチャートである。同図に示される如く、本発明に係る顔画像認識システムでは、最初に説明したEigen-Window法に基づくエッジ強度に基づく窓選択に先立って、肌色情報に基づく窓選択を行う(S1101)。認識対象が人間の顔画像であることに基づくEigen-Window法の改良点の一つである。

【0061】肌色情報に基づく窓選択とは、初期設定処理において、本実施の形態の顔画像認識システムに認知された肌色情報に基づき、学習画像において肌色に相当する領域のみから窓選択を行う処理である。このような

10

20

30

40

50

処理を行うことにより、エッジ強度に基づく窓選択等、その後の処理の効率を向上させることが可能となる。

【0062】図12は、肌色情報に基づく窓選択の詳細な処理内容を示すフローチャートである。同図に示されるように、肌色情報に基づく窓選択においては、まず、各学習画像にH S I変換を施し、変換された各画素のH S I値を基に、人間の顔の肌色に相当する領域を抽出する(S1201)。

【0063】上述の如く、初期設定処理において肌色領域に該当するH S I値の範囲が抽出されているので、ここでは、当該肌色領域に該当するHue値hを示す画素のみを選択することになる。

【0064】次に、一致度マップの生成を行う(S1202)。一致度マップとは、上記ステップS1201にて選択された画素のH S I値について、前記初期設定処理にて選択された肌色領域内における位置から計算される一致度の値を画像上の画素位置に書き込んだものである。一致度の計算は、具体的には以下のようにして行う。即ち、例えば、肌色領域の直方体の中心に近い位置にプロットされた点に対しては高い一致度を割り当て、直方体の側面に近い位置にプロットされた点に対しては低い一致度を与える。なお、肌色領域の外部に位置する点に対しては一致度0を割り当てる。図13(a)は、一致度マップの内容を具体的に表した図である。

【0065】さらに、この一致度マップに対して、閾値処理を行い、あらかじめ設定された閾値よりも高い一致度を示す画素を窓画像候補点として選択する(S1203)。図13(b)に、選択された窓画像候補点の内容を具体的に示す。

【0066】上記、肌色情報に基づく窓選択を終了すると、図11のフローチャートに戻って、エッジ強度に基づく窓選択を行う(S1102)。エッジ強度に基づく窓選択については、上記Eigen-Window法の説明において既に説明しているので、ここでの詳細な説明は省略するが、本実施の形態の顔画像認識システムにおいては、上記肌色情報に基づく窓選択の処理にて選択された肌色を示す画素に対してのみ、画像の濃淡値を基にエッジ強度が計算され、窓選択が行われることになる。

【0067】次に、本実施の形態の窓画像認識システムでは、明るさの正規化処理を行う(S1103)。これは、顔画像の撮影時に起こりうる照明及び環境の変化等に起因する対象の見えの変化に対して、ある程度ロバストな認識を可能にするためにEigen-Window法を改良した点の一つである。

【0068】図14は、明るさの正規化処理の詳細な処理内容を示すフローチャートである。同図に示されるように、明るさの正規化処理においては、まず、窓画像として選択されたR G B値に対してGrav-Scale変換を施し、一つのGrav-Scale値(通常、256段階の濃淡値である。)で表現されたGrav-S

cale窓画像に変換する(S1401)。

【0069】次に、この窓画像の各画素のGrav-Scale値 $z_1, z_2, \dots, z_N$ (Nは窓画像の全画素数)を要素とする窓画像ベクトルZを生成し(S1402)、このベクトルの各要素 $z_i$ の2乗和のルート(以下、この値をベクトルの「ノルム」という。)を計算する(S1403)。以下、窓画像ベクトルZのノルムの値を $\|Z\|$ で表す。

【0070】さらに、窓画像ベクトルZをこのノルムの値 $\|Z\|$ で除算することにより、明るさの正規化処理を行い、その結果得られた正規化窓画像ベクトルZ'を、明るさの正規化処理の結果として出力する(S1404)。

【0071】出力された正規化窓画像ベクトルZ'について、KL展開を行うことによって(S1105)、学習窓画像対応点の情報がモデル画像格納部103に格納されるが(S1106)、KL展開については、Eigen-Window法に関する説明のところで詳述しているので、ここでの詳細な説明は省略する。

【0072】尚、KL展開を行うためには、準備された全ての学習画像を読み込んでおく必要があるので、KL展開を行う前に、学習画像をすべて読み込むようにしている(S1104)。

【0073】次に、準備された全ての学習画像をKL展開し、その結果をモデル画像格納部103に格納すると、学習フェーズが終了する。

【0074】次に、モデル画像格納部103に格納された学習窓画像対応点の情報を用いて、認識対象画像を認識する認識フェーズにおける、本実施の形態の顔画像認識システムの処理内容について説明する。

【0075】図15は、認識フェーズにおける認識処理部1022の処理内容を示すフローチャートである。同図に示される如く、認識フェーズでは、認識対象となる顔画像情報について、学習フェーズと同様に、肌色情報に基づく窓選択(S1501)、エッジ強度に基づく窓選択(S1502)、明るさの正規化(S1503)を行う。それぞれの処理の内容については、学習フェーズで説明した内容と同一であるので、ここでの詳しい説明は省略する。

【0076】認識フェーズでは、明るさの正規化を行った後、KL展開(S1504)、最近傍点の探索(S1505)、及び投票処理(S1506)を行う。

【0077】この部分の処理内容については、最初にEigen-Window法について説明したところで詳述しているので、ここでの説明は省略する。

【0078】以上のような、認識フェーズの処理を行うことにより、入力された顔画像に対応する人物の名前、姿勢等を特定することができる。

【0079】

【実施例】図16に、本発明の一実施例における顔画像

10

20

30

40

50

認識システムの構成を示す。

【0080】同図に示されるように、本実施例においては、本発明の顔画像認識システムは、床に固定された椅子1601、椅子の角度を計測するためのゲージ1602、顔画像撮影用カメラ1603、および認識処理を行うコンピュータ1604などから構成されている。

【0081】学習した範囲内の任意の角度を向く椅子に腰掛けた人物の顔画像をカメラから取り込み、この入力された顔画像と、あらかじめ学習画像から作成したモデルとの照合を行うことにより、顔の識別と姿勢同定を行うことができる。

【0082】まず、学習フェーズでは、学習画像として、5人の人物を撮影用カメラから1250mm離れた椅子に座らせて、0度（右真横）から180度（左真横）まで30度おきに撮影した計35枚の画像を用いた。図17は、本実施例において使用した学習画像を示す図である。

【0083】撮影時には特別な照明は使用せず、室内灯のみの環境下で行った。

【0084】取得した画像のサイズは120×160画素であり、そのRGB値を濃淡値画像に変換した後、エッジ強度に基づく窓画像（N=10）選択を行い、k=10次元の固有空間上へ投影した。

【0085】この時の35枚の学習画像に対する総窓画像数は3485個であった。

【0086】次に認識フェーズでは、以下に示す三つの場合のそれぞれに対して、5人の人物を0度から180度まで15度おきに撮影し、これら計195枚の画像（学習画像と同サイズ）を評価用の入力画像（認識対象画像）として用いた。図18に、本実施例の認識フェーズで使用した入力画像を示す。

【0087】第1の実施例として、カメラ設置位置を120mm移動し、相対的に画像中の顔位置を移動させた画像を入力した場合（65枚）、第2の実施例として、被撮影者が口元を手で覆い隠した画像を入力した場合（65枚）、及び第3の実施例として、学習画像と異なる複雑な背景を含む画像を入力した場合（65枚）のそれぞれについての、実施の結果を図19および図20に示す。

【0088】図19に示された数字は、顔の姿勢に関わらず、入力画像中に含まれる人物を正しく特定できた場合を正解とした個人識別率を表している。また、図20に示された数字は、顔の姿及び入力画像中に含まれる人物を正しく特定できた場合を正解とした個人識別／姿勢特定率を表している。

【0089】尚、いずれの場合も、比較のため、固有空間法を用いた場合の個人識別率を下段に示している。

【0090】このように、本発明のに係る顔画像認識システムは、従来の固有空間法に基づく同様のシステムと比較して、対象の位置ずれ、隠れ、背景ノイズが存在す

る場合においても、識別率が向上することが確認された。

【0091】

【発明の効果】以上に説明したように、本発明に係る顔画像認識システムは、従来の固有空間法が有していた、位置ずれに弱い等の欠点を克服したEigen-Window法に基づく顔画像認識手法を用いるため、認識対象画像が対象の位置ずれや隠れ、背景ノイズを含む場合、また複数の顔が画像内に存在する場合などにも対応することが可能な顔画像認識システムを実現することができるという効果を奏する。

【0092】また、HSI変換した後のHue値を用いて、顔に相当する肌色領域に対してのみから窓画像を選択するため、顔の部分以外の部分に位置する窓画像に対する処理を極力排除でき、結果的に、外乱を含む画像に対しても認識率を向上させるとともに、認識フェーズにおける処理速度を向上させることができるという効果がある。

【0093】さらに、各窓画像ごとに明るさの正規化処理を施すため、画像全体の明るさが均一に変動した場合だけでなく、影などにより局所的に明るさが変動する場合においても、安定な顔画像認識が可能になるという効果がある。

【図面の簡単な説明】

【図1】 本発明の一実施の形態に係る顔画像認識システムの全体構成を示すブロック図である。

【図2】 Eigen-Window法の学習フェーズにおける処理過程の概略を示す図である。

【図3】 窓選択について説明するための図である。

【図4】 学習窓画像対応点の様子を表す図である。

【図5】 Eigen-Window法の認識フェーズにおける処理過程の概略を示す図である。

【図6】 最近傍点の探索処理における固有空間の様子を具体的に示す図である。

【図7】 投票空間の様子を具体的に示す図である。

【図8】 初期設定処理における認識処理部の処理内容を示すフローチャートである。

【図9】 初期設定処理において用いられるサンプル画像の一例を示す図である。

【図10】 HSI空間における肌色領域の抽出について、具体的に説明するための図である。

【図11】 学習フェーズにおける認識処理部の処理内容を示すフローチャートである。

【図12】 肌色情報に基づく窓選択の詳細な処理内容を示すフローチャートである。

【図13】 (a) 一致度マップの内容を具体的に説明するための図である。

(b) 選択された窓画像候補点の内容を具体的に示す図である。

【図14】 明るさの正規化処理の詳細な処理内容を示

すフローチャートである。

【図 15】 認識フェーズにおける認識処理部の処理内容を示すフローチャートである。

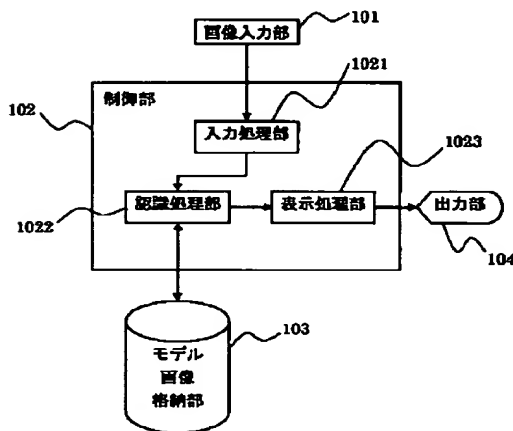
【図 16】 本発明の一実施例における顔画像認識システムの構成を示す図である。

【図 17】 本実施例の学習フェーズにおいて使用した学習画像を示す図である。

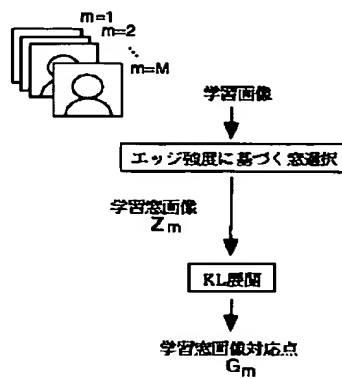
【図 18】 本実施例の認識フェーズにおいて使用した入力画像を示す図である。

【図 19】 本実施例において、顔の姿勢に関わらず、10  
入力画像中に含まれる人物を正しく特定できた場合を正解とした個人識別率を表す図である。 \*

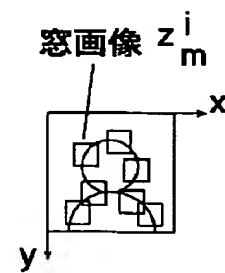
【図 1】



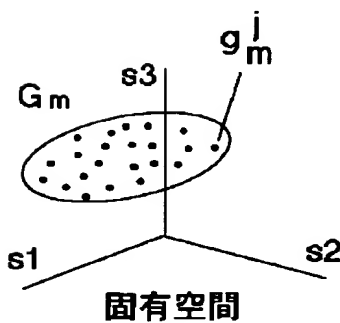
【図 2】



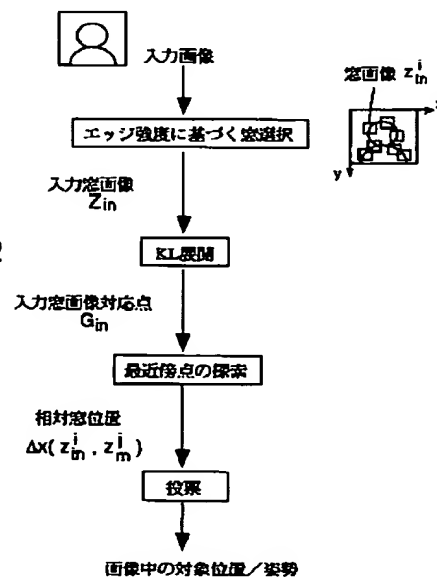
【図 3】



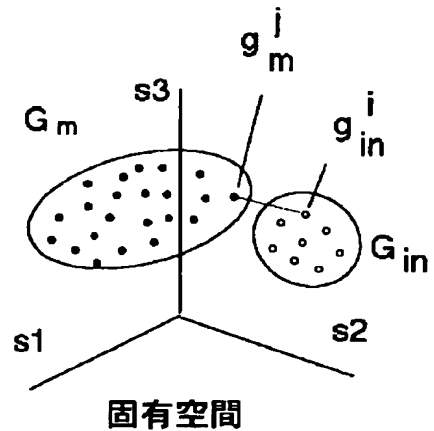
【図 4】



【図 5】



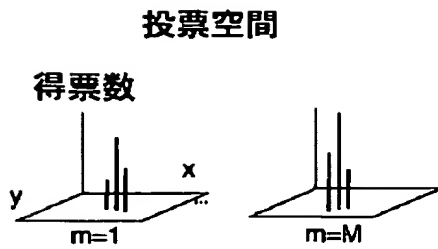
【図 6】



【図 18】



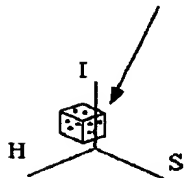
【図 7】



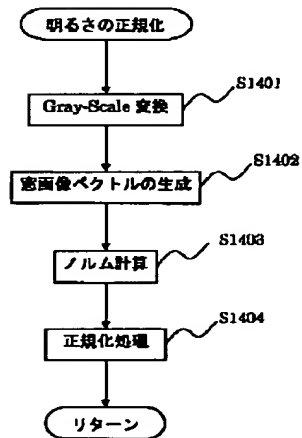
【図 10】

## 肌色領域

$$\begin{cases} H_{min} < h < H_{max} \\ S_{min} < s < S_{max} \\ I_{min} < i < I_{max} \end{cases}$$



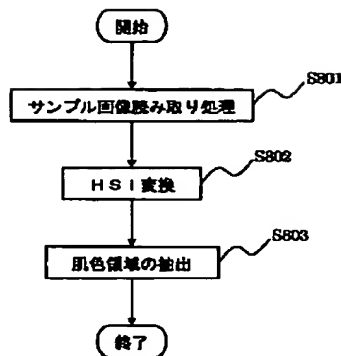
【図 14】



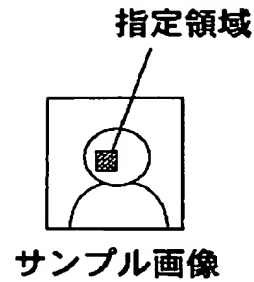
【図 19】

|       | Case 1 | Case 2 | Case 3 |
|-------|--------|--------|--------|
| 本発明   | 95.4%  | 90.8%  | 98.5%  |
| 固有空間法 | 29.2%  | 84.8%  | 67.7%  |

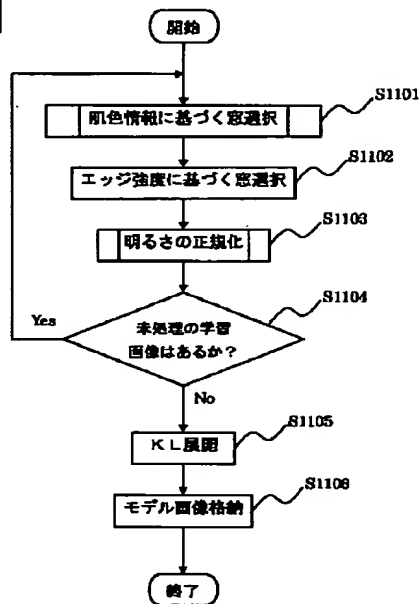
【図 8】



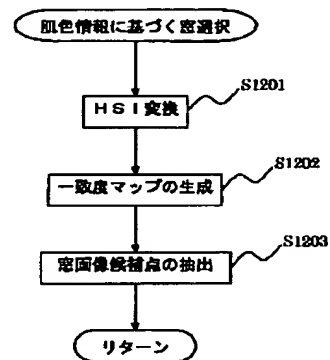
【図 9】



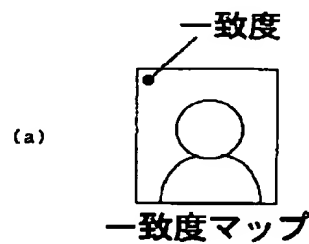
【図 11】



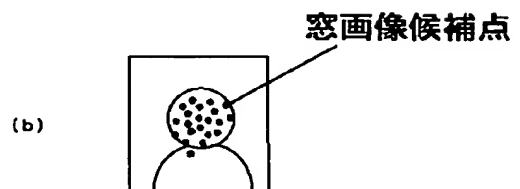
【図 12】



【図 13】

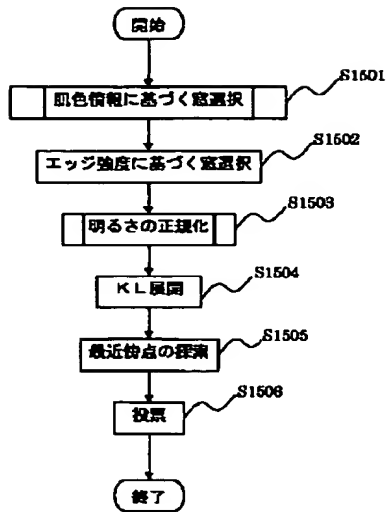


(a)

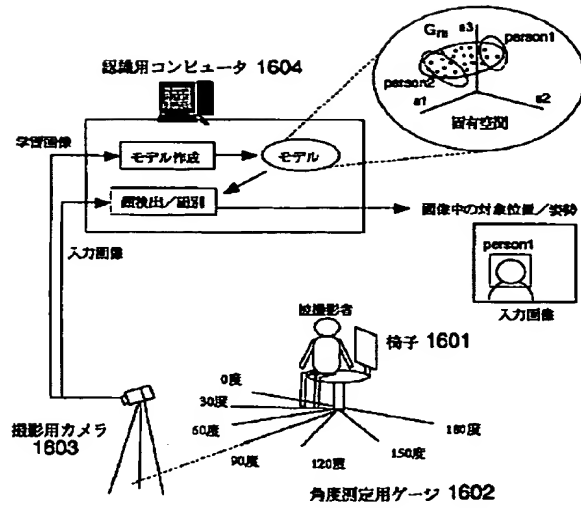


(b)

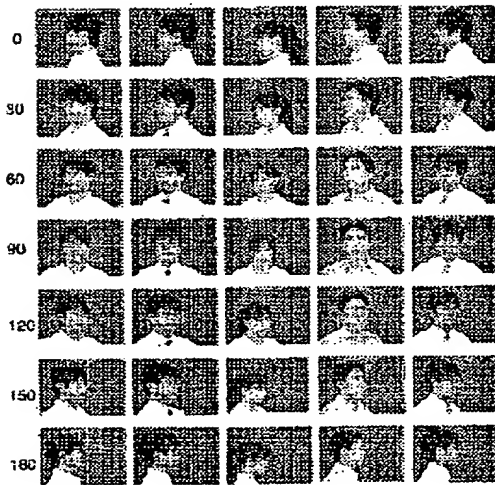
【図 15】



【図 16】



【図 17】



【図 20】

|       | Case 1 | Case 2 | Case 3 |
|-------|--------|--------|--------|
| 本発明   | 84.5%  | 87.7%  | 88.5%  |
| 面有空間法 | 8.1%   | 83.1%  | 48.2%  |